UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/608,590	06/27/2003	James M. Sweet	D/A2555Q	8455

77282 7590 06/11/2009

Hoffmann & Baron, LLP
6900 Jericho Turnpike
Syosset, NY 11791

EXAMINER	
HILLERY, NATHAN	
ART UNIT	PAPER NUMBER
2176	

MAIL DATE	DELIVERY MODE
06/11/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 10/608,590
Filing Date: June 27, 2003
Appellant(s): SWEET ET AL.

Christopher D. Wait
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 10/10/08 appealing from the Office action mailed 12/26/07.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The following are the related appeals, interferences, and judicial proceedings known to the examiner which may be related to, directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal:

10/608,587

10/608,591

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is substantially correct. The changes are as follows:

GROUNDS OF REJECTION NOT ON REVIEW

The following grounds of rejection have not been withdrawn by the examiner, but they are not under review on appeal because they have not been presented for review in the appellant's brief.

Claims 21 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Huang et al. (US 6601075 B1).

Claims 23 and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Law et al. (US 6754873 B1).

Claims 7 – 9, 11, 19, 20, 24, 31 – 34, and 45 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Prince (US 6877002 B2).

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

6,112,203	Bharat et al.	8-2000
5,924,104	Earl	7-1999
6601075	Huang et al.	7/2003
6754873	Law et al.	6/2004
6877002	Prince	4/2005

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and in further view of Earl (US 5924104 A).

Regarding independent claim 37, Bharat et al. teach that we locate pages that point to at least one of the pages in the start set 201. We call this set of pages the back set 202 (Column 4, line 61 – Column 5, line 20), which meets the limitation of **performing a page-level link analysis that identifies those hyperlinks on a page linking to a candidate document page**.

Bharat et al. teach that if a link points to a page that is represented by a node in the graph, and both pages are on different servers, then a corresponding edge 213 is added to the graph 211. Nodes representing pages on the same server are not linked. This prevents a single Web site with many self-referencing pages to unduly influence the outcome (Column 4, line 61 – Column 5, line 20), which meets the limitation of **identifying possible progression links**; **identifying possible table of content links**; and **examining the possible progression links and the possible table of content**

links for common characteristics, since the specification states that "the possible progression links 230 and possible table of content links 240 are passed to module 250 for a final examination to weed out links which have properties that are not characteristic of typical intra-document links, e.g. they point to a different web server" (p 7, lines 26 – 30). It should be noted that pages on the same server are nodes and are thus still apart of the resulting graph.

Furthermore, it should be noted that the self-referencing pages of Bharat et al. are equivalent to intra-document links and that those intra-document links can be "possible" progression and/or table of contents links, since the Office has interpreted the word "possible" as "could be" and within the broadest, reasonable interpretation in light of the specification, which states that a link analysis phase consists of the identification for a given hypertext page of the most likely desirable intra-document links. Those intra-document links fall into two categories: progression links and table of contents links (p 5, second paragraph). Thus, any intra-document link – a link that points to a different web server – could be a possible progression or table of contents link.

Bharat et al. teach that a larger n-graph 211 can be constructed by repeating this process for the back and forward sets 202-203 to add more indirectly linked pages (Column 4, line 61 – Column 5, line 20), which meets the limitation of **performing a recursive application of the page-level link analysis to the linked candidate document page and any further nested candidate document pages thereby identified, until a collective set of identified candidate document pages is assembled**.

Bharat et al. do not explicitly teach **performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents; examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of typical intra-document links, to provide a resultant set of identified candidate document pages; and grouping the content found in the resultant set of candidate document pages into a document representation for subsequent viewing or printing of the given hyperdocument**.

Earl teaches that the link display manager 300 includes a document parser 304 for parsing each document and identifying links 202 and 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **performing a document-level analysis that examines the collective set of identified candidate document pages for grouping into one or more documents**.

Earl teaches that the link display manager 300 includes a display system for defining predetermined screen element properties providing visual cues for distinguishing the identified links 202 and 204. When a user provides an input link selection to select a new document, the document parser 304 parses the selected new document to identify intradocument links 202 and interdocument links 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages**.

Earl teaches that the display system 306 processes the identified intradocument links 202 and interdocument links 204 for displaying distinctively the intradocument links 202 and interdocument links 204 with predetermined visual cues to differentiate the links 202, 204 (Column 2, line 59 – Column 3, line 9), which meets the limitation of **grouping the content found in the resultant set of candidate document pages by an automated system into a document representation stored in memory by the automated system and printing or viewing on a display by a user the document representation**.

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the invention of Bharat et al. with that of Earl because such a combination would provide the users of Bharat et al. with *an improved method and apparatus for displaying links on a user display interface in a computer system* (Column 1, lines 39 – 41).

Regarding dependent claim 38, Bharat et al. teach that the nodes in the start set are first scored according to their connectivity, and the number of terms of the query that appear as unique sub-strings in the URL of the represented documents. The score is a weighted sum of the number of directed edges to and from a node and the number of unique sub-strings of the URL that match a query term (Column 3, lines 10 – 15), meets the limitation of **the page-level link analysis includes examination of contextual clues**, **the contextual clue is a particular class of content item**

associated with the hyperlink, **the class of content item is a class of text**, **the class of text is a directional word or phrase**.

Regarding dependent claim 46, Bharat et al. teach that we assign a similarity weight to each node 213 of the sub-graph 255. Various document similarity measuring techniques have been developed in Information Retrieval to determine the goodness of fit between a "target" document and a collection of documents. These techniques typically measure a similarity score based on word frequencies in the collection and a target document (Column 6, lines 51 – 57), meets the limitation of **the contextual clue is the similarity of the hyperlink destination to that of other hyperlinks within the hyperdocument**.

Regarding dependent claim 47, Bharat et al. teach that we use a modified Kleinberg algorithm on the nodes of the pruned n-graph 265 to determine useful hub and authority pages. For each node of the pruned n-graph 265, we measure two scores: a hub score (HS), which estimates how good a hub the page is, and an authority score (AS), which estimates how good an authority the page is. The intuition behind our method is this: a good hub is one that points to many documents. A good authority is one that is pointed to by many documents. Transitively, an even better hub is one that points to many good authorities, and an even better authority is one that is pointed to by many good hubs (Column 7, lines 41 – 50), meets the limitation of **the**

document-level analysis includes the identification of pages forming a chain of progression links.

Regarding dependent claims 18 and 48, Bharat et al. teach that after we have constructed the nodes 212, we add the directed edges 213. If a link points to a page that is represented by a node in the graph, and both pages are on different servers, then a corresponding edge 213 is added to the graph 211. Nodes representing pages on the same server are not linked. This prevents a single Web site with many self-referencing pages to unduly influence the outcome. This completes the n-graph 211 (Column 5, lines 13 – 20), meets the limitation of **the similarity includes the location at which the page is stored**, and **the document-level analysis includes the identification of pages linked to by the same tables of contents**.

Regarding claims 1 – 6, 10 and 12 – 14, the claims incorporate substantially similar subject matter as claims 37, 38 and 46 – 48 and are rejected along the same rationale.

Regarding claims 26 – 30, 35 and 36, the claims incorporate substantially similar subject matter as claims 37, 38 and 46 – 48 and are rejected along the same rationale.

Regarding dependent claim 15 – 17, Bharat et al. teach that we use do iterative connectivity analysis 310, content analysis 320, and pruning 330. This method consists of a sequence of rounds. In each round, our modified connectivity analysis is run for 10 iterations to get a listing of the (current) best hubs and authorities 315. In step 320, the pages are examined for content similarity in decreasing order of rank, alternating between the hub and the authority list. Less relevant pages are pruned (Column 8, lines 25 – 33), meets the limitation of **the document-level analysis includes identifying the pages listed in a table of contents**, **the document-level analysis includes identifying as part of the document the page containing the table of contents**, **the document-level analysis includes the similarity of candidate pages**.

Claims 21 and 22 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Huang et al. (US 6601075 B1).

Regarding dependent claims 21 and 22, neither Bharat et al. nor Earl teach **the similarity includes similar style specifications**, and **the similarity includes similar page layout.**

Huang et al. teach that the HITS and CLEVER algorithms make use of hyperlinked structures to rank documents that share the same schema. Exemplary documents with hyperlinked structures are HTML documents. XML has given rise to a new hyperlink environment that includes documents with different schemas. Hence, this

new environment helps rank documents based on the quality of their associated schema, determine the quality of the schemas themselves, and ranks documents based on their structural properties (e.g. validity, well-formedness, etc.) (Column 3, lines 37 – 53), meets the limitation of **the similarity includes similar style specifications**, and **the similarity includes similar page layout**.

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the combined invention of Bharat et al. and Earl with that of Huang et al. because such a combination would allow the users of Bharat et al. and Earl the benefit of *an algorithm which is applied to an initial set of documents, similar to the HITS and CLEVER algorithms* (Column 3, line 66 – Column 4, line 1).

Claims 23 and 25 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Law et al. (US 6754873 B1).

Regarding dependent claims 23 and 25, neither Bharat et al. nor Earl teach **the similarity includes similar logical structure of the page content**, **the document-level analysis includes analysis of the topological structure of the linked pages**.

Law et al. teach that the link structure of the hyperlinked documents is analyzed in order to find hyperlinked documents that are related to and at the same level of generality of a hyperlinked document (Column 2, lines 8 – 11), meets the limitation of

the similarity includes similar logical structure of the page content, **the document-level analysis includes analysis of the topological structure of the linked pages**.

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the combined invention of Bharat et al. and Earl with that of Law et al. because such a combination would allow the users of Bharat et al. and Earl the benefit of *innovative techniques for finding related hyperlinked documents using link-based analysis* (Column 2, lines 6 – 8).

Claims 7 – 9, 11, 19, 20, 24, 31 – 34, and 45 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A) and Earl (US 5924104 A) as applied to claims 1 – 6, 10, 12 – 18, 26 – 30, 35 – 38, and 46 – 48 above, and further in view of Prince (US 6877002 B2).

Regarding dependent claims 7 – 9, 11, 19, 20, 24, 31 – 34, and 45, neither Bharat et al. nor Earl explicitly teach **meta-data** or **image**.

However, Prince teaches that the parsed results (from step 42 in FIG. 4) relating to the media are passed to extraction agent 68 via an extraction queue 67. The extraction queue 67 comprises URLs to be analyzed with respect to associated media metadata. The extraction queue 67 may comprise metadata queue entries such as media URLs, Web page URLs, Web page titles, Web page keywords, Web page descriptions, media title, media author, and media genre. Each queue entry added to the extraction queue is assigned a processing time and a priority (Column 7, lines 23 –

37), meets the limitation of **the similarity includes the similarity of meta-data associated with the page**, **the meta-data includes the author identification, the similarity includes the presence of at least one similar content item on each page**, **the class of content item is a class of image**, **the class of image is an image containing a directional symbol**, **a textual clue is obtained for the class of image**, **the contextual clue is the presence of at least one other hyperlink nearby with the candidate document page**.

It would have been obvious to one of ordinary skill in the art at the time of the invention to combine the combined invention of Bharat et al. and Earl with that of Prince because such a combination would allow the users of Bharat et al. and Earl the benefit of *A method for querying metadata associated with media on a computer network includes separating the metadata into keywords* (Column 2, lines 37 – 39).

(10) Response to Argument

For the sake of brevity, Appellant attacks Bharat for not teaching **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages** because Bharat is only interested in inter-document links (pp 9 – 12).

The Office does not know what to make of these arguments because Bharat is not relied upon to teach this limitation or intra-document links. Thus, these arguments appear to be moot.

Appellant argues that Earl does not teach **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages** because what Earl defines as intra-document is what Appellant would call intra-page and thus what Earl calls inter-document is really inter-page (pp 12 and 13).

The Office disagrees.

First, it should be noted that the claim recites "links which have properties that are not characteristic of intra-document links," which has been interpreted as links that appear, seem, or are otherwise similar to intra-document links. Regardless of how appellant would like the Office to interpret the intra-document links of Earl versus those defined by Appellant's specification, Appellant fails to positively recite intra-document links and chooses only to recite links that may be characteristic of intra-document links. Thus within the broadest, reasonable interpretation in light of the specification, the Office maintains that the intra-document links or inter-document links of Earl read on or renders obvious the links which have properties that are not characteristic of intra-document links.

Further, the claim's use of double negatives and open language lends itself to being interpreted loosely with bigger metes and bounds that appellant may intend. As to the declarations, the declarations raise the question of how much weight they hold. They appear to be self-serving and give the impression that one must just take

appellant's word for it. The audacity to imply that the Earl reference is misusing terms of art such as intra-document and inter-document links and to attempt to explain that the Earl reference really mean circular links is appalling and unconvincing.

In light of all the evidence, the claim recites **links, which have properties that are *not characteristic* of intra-document links**. Consequently, the Office maintains that Earl clearly and explicitly teaches intra-document and inter-document links that meet the claimed intra-document link, since both or at least one of the links, intra-document and inter-document, have properties that are characteristic of intra-document links within the broadest, reasonable interpretation in light of the Specification.

Appellant argues that Earl further does not teach **examining the collective set of identified candidate document pages to weed out links which have properties that are not characteristic of intra-document links, to provide a resultant set of identified candidate document pages** because Earl keeps all the links while appellant's invention discards or "weeds" out the undesirable links (pp 13 and 14).

The Office disagrees.

The Office maintains that Earl does "weed out" the links within the broadest, reasonable interpretation in light of the specification. The Declaration, which appellant cites to bolster this argument, actually states that "in the specification, the appellants used the phrase 'filtered out' as an alternative to the 'weed out' language, which will also further reinforce for those skilled in the art as to what the appellants mean" ('Sweet' Declaration, p 4, item 9).

As it relates to computers, the term 'weed out' does not necessarily require a delete operation especially in light of the fact that appellant means 'weed out' to be synonymous with 'filter out.' Filtered out is a term of art that basically means to discriminate or distinguish. By design, computers are meant to hold and store information. Just because at time t1, the user would like certain data weeded out does not mean that at time t2, the user has no use for that data and would not want it then. Computers allow a user to hone in on certain data at certain times depending on what data the user would like to focus at that time; permanently deleting a subset of data for one operation automatically would cause catastrophic results for most users and would be counterintuitive to the skilled artisan.

Thus, Earl, by appellant's own admission, teaches discriminating visually between intra-document and inter-document links (p 13, first full paragraph), which meet the definition of separating out, weeding out, or filtering out, the links visually on screen. The requirement to have to discard the links is still too limiting in view of what is actually claimed and in light of all the evidence.

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

/Nathan Hillery/

Examiner, Art Unit 2176

Conferees:

/Doug Hutton/
Doug Hutton
Supervisory Primary Examiner
Technology Center 2100

/Rachna S Desai/
Rachna S. Desai
Primary Examiner, Art Unit 2176